

Where To Go Next?

Exploiting Behavioral User Models in Smart Environments

Seyyed Hadi Hashemi
University of Amsterdam
Amsterdam, The Netherlands
hashemi@uva.nl

Jaap Kamps
University of Amsterdam
Amsterdam, The Netherlands
kamps@uva.nl

ABSTRACT

There is a growing interest in using the Internet of Things (IoT) to create smart environments, which hold the promise to provide personalized experience based on the trail of user interactions with smart devices. We experiment with behavioral user models based on interactions with smart devices in a museum, and investigate the personalized recommendation of what to see after visiting an initial set of Point of Interests (POIs), a key problem in personalizing museum visits or tour guides. We have logged users' onsite physical information interactions of visits in a museum. Moreover, to have a better understanding of users' information interaction behaviors and their preferences, we have collected and studied query logs of a search engine of the same collection, and we have found similarities between users' online digital and onsite physical information interaction behaviors. We exploit user modeling based on users' different information interaction behaviors and experiment with a novel approach to a critical one-shot POI recommendation using deep neural multilayer perceptron based on explicitly given users' contextual information, and set-based extracted features using users' physical information interaction behaviors and similar users' digital information interaction behaviors. Experimental results indicates that our proposed behavioral user modeling, using both physical and digital user information interaction behaviors, improves the onsite POI recommendation baselines' performances in all common Information Retrieval evaluation metrics. Our proposed approach provides an effective way to achieve a high precision at rank 1 in onsite critical one-shot POI recommendation problem.

KEYWORDS

Human information interaction, Onsite logs, Behavioral user models, POI recommendation, Internet of things

ACM Reference format:

Seyyed Hadi Hashemi and Jaap Kamps. 2017. Where To Go Next? Exploiting Behavioral User Models in Smart Environments. In *Proceedings of UMAP'17, July 9-12, 2017, Bratislava, Slovakia*, 9 pages. DOI: <http://dx.doi.org/10.1145/3079628.3079687>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

UMAP'17, July 9-12, 2017, Bratislava, Slovakia

© 2017 ACM. 978-1-4503-4635-1/17/07...\$15.00

DOI: <http://dx.doi.org/10.1145/3079628.3079687>



Figure 1: Interactive POIs in a museum physical space

1 INTRODUCTION

The last decade witnessed a tremendous interests in implementation of Internet of Things (IoT) in different applications[2, 3, 10, 16–18, 25], such as smart shopping malls and smart museums, which provides the infrastructure for understanding users' physical interaction behavior and consequently their preferences in interacting with smart environments. This prompts the question:

How tracking people in their real-life and understanding their interaction behaviors would be helpful? Is it possible to give effective suggestions to users by user tracking using IoT but without getting any explicit information about their preferences like ratings?

Imagine you are at a huge museum like the Louvre Museum in Paris and you want to explore the museum. Usually, it is impossible to visit the whole objects of some big museums like the Louvre Museum in one day. Moreover, museum free roaming is more desirable in comparison to the traditional fixed walking route designed in a non-personalized way. Providing personalized experiences for users is so valuable in this context and will help them to visit all interesting objects of the museum according to their preferences. In this case, how amazing would it be if the contextual recommender system can tell you accurately what to visit without using any history or explicit feedback from you?

Emergence of the above applications leads to rise interests in logging users' onsite physical information interactions, which creates a new potentially exponentially growing data like search engine query logs. Although understanding users' search behavior and their information needs based on query logs is well-studied [7, 29], to the best of our knowledge, there is not any study on how to understand users' behaviors and their information needs based on similarities between users' onsite physical and online digital information interaction behaviors. Addressing this research problem by

2 RELATED WORK

In this section, we discuss related work on context aware recommendation systems, POI recommendation systems, and recommendation systems in museums.

Traditionally, recommendation systems deal with applications having just two types of entities, users and items. However, creation of more complex and realistic applications leads to interests in a new line of research about how to incorporate contextual information as an extra dimension to the recommendation systems. There are 3 ways of incorporating context in the recommender systems: contextual pre-filtering, contextual post-filtering, and contextual modeling [1]. As the later approach is closer to our study in this paper, we will discuss some of the related research in the contextual modeling.

In order to contextually model the context aware recommendation system, Karatzoglou et al. proposed a multiverse recommendation method based on tensor factorization [21], which integrate contextual information by modeling data as User-Item-Context N-dimensional tensor instead of traditional 2-dimensional User-Item matrix. One problem of this method is the data sparseness, which is proportional to the number of defined context in their method. Liu et al. [22] proposed to partition the User-Item matrix by grouping ratings of similar context, which could be helpful to decrease the data sparseness. The other problem of the multiverse recommendation method is that it only works for categorical features. To overcome this problem, Rendle et al. [26] proposed to use factorization machines to model contextual information. The above studies are done to model contextual information, however none of them are really scalable and effective for the recent exponentially growing data.

There have also been lots of studies to solve the POI recommendation problem in both academia and industry [14, 38]. They generally try to adapt traditional recommendation algorithms to the POI recommendation problem. One line of research includes collaborative filtering and matrix factorization approaches in location-based social networks (LBSNs). Berjani et al. in [5] proposed regularized matrix factorization, in which they apply personalized collaborative filtering on dimensionally reduced user-POI matrices to minimize the squared regularized error. In addition to the geographical aspects, they are some researches in POI recommendation that in addition to the geographical dimension, tried to include temporal dimension in the matrix factorization framework [11, 13].

Within the POI recommendation literature, there are some studies that are related to ours in the sense that they studied users' check-in behavior [24, 27, 30–34, 36, 37]. As three interesting examples of these related works, Zheng et al. proposed collaborative location activity filtering [35]. Particularly, they used collective factorization to recommend locations or activities to users. To this aim, they used comments having GPS data in a web-based GPS management system as a data source. Moreover, Ye et al. in [31] proposed a collaborative POI recommendation algorithm based on geographical influence. To this aim, they used users check in activities in LBSNs. At last, Scholz et al. [27] studied talk attendance prediction in an academic conference using a link prediction approach. To this aim, they logged talk attendance behavior using RFID tags. However, none of the above studies used both the actual

users' onsite physical information interaction behaviors and users' online digital click-through behaviors.

As another line of related research, there are several researches that study recommender systems for museum visitors. Grieser et al. [12] studied next exhibition recommendation problem in the museum space using visitors history. They applied Naive Bayes learning model using textual description, geospatial proximity and popularity of exhibitions. In their study, popularity baseline, which is one of our defined baseline in this paper, was reported as the most successful next exhibition recommendation model.

Bohnert et al. [6] studied unseen exhibition recommendation using nearest-neighbor content-based filtering approach by taking visitors explicit ratings of exhibitions as inputs. They did the study using 41 museum visitors as participants. Moreover, in a recent work of Bartolini et al. [4], they study recommendation of diverse multimedia data across several web repositories, and arrangement of them in visiting paths. They consider location, number of persons and weather condition as context in their contextual pre-filtering system, and they did the study based on 90 users as participants.

Apart from different recommendation methods being used in the above studies in the museum domain, they are very limited in term of number of participants. In addition, none of them log and study users' onsite physical information interactions behaviors. In this paper, we log more than 21,000 users' visits of a museum in a 5 months period, and our proposed model is based on users' both online digital and onsite physical information interaction behaviors.

Closest in spirit to our work is [17], in which users' onsite physical behaviors in the existence of a crowd of users have been studied. They studied skip or stay behavior prediction in checking in different POIs as a classification problem. Their study is different from ours as they do not investigate on similarities between users' physical and digital behaviors. Moreover, we study a POI ranking problem in this paper but they did research on onsite physical interaction behavior classification problem.

3 BACKGROUND AND PRELIMINARIES

In this section, we state the behavioral unseen POI recommendation problem and the best baselines comparable with our proposed model.

3.1 Problem Statement

Let $\mathbf{u} = \{u_1, u_2, \dots, u_i\} \subset U^i$ be a subset of users visited a smart environment, $\mathbf{c}_{seen} = \{c_1, c_2, \dots, c_j\} \subset C_{seen}^j$ a subset of seen or occurred contexts, and $\mathbf{p}_{seen} = \{p_1, p_2, \dots, p_k\} \subset P_{seen}^k$ a subset of seen POIs. Then, let $\mathbf{R}_{seen} \in \mathbb{R}_{seen}^{i \times j \times k}$ be a user-context-POI matrix containing i users, j seen contexts and k seen POIs. Value $r_{i,j,k} \in \mathbf{R}_{seen}$ refers to the visit frequency of user i , in context j to the POI k . In this paper, due to the fact that museum visitors rarely check in to a POI more than once, we have used binary seen or unseen values rather than considering the frequency.

Having above information about users, given a subset of unseen contexts (i.e., $\mathbf{c}_{unseen} = \{c_1, c_2, \dots, c_m\} \subset C_{unseen}^m$), and a subset of unseen POIs (i.e., $\mathbf{p}_{unseen} = \{p_1, p_2, \dots, p_n\} \subset P_{unseen}^n$), the behavioral unseen POI recommendation problem is estimation of $r_{i,m,n} \in \mathbf{R}_{unseen}$ based on users interaction behaviors with the

seen POIs, in which $\mathbf{R}_{unseen} \in \mathbb{R}_{unseen}^{i \times m \times n}$ is a user-context-POI matrix containing i users, m unseen contexts and n unseen POIs.

3.2 Baselines

In this section, baselines created for the evaluation purposes are detailed.

3.2.1 Popularity. The popularity based recommendation ranks POIs candidates according to their popularity scores. The popularity is computed as the number of users who checked in at each POI. The popularity baseline is usually used in evaluation of personalized recommendation systems and it is informed as a very challenging and hard-to-beat baseline [23].

3.2.2 Bias-Based Filtering. As Hashemi et.al. discussed in [16], there are some biases in onsite user information interaction logs. They introduces the walk-through position-bias that shows users tend to visit POIs one after the other from check-in to check-out stations. They also observed time-rank bias that indicates users tend to spend less time at the end of exhibitions. Considering these two biases, the probability of checking in at a POI is proportional to the distance from the Check-out station. Therefore, Bias-based baseline ranks POIs based on their distance from the check-out station.

3.2.3 Content-Based Filtering. As descriptions of POIs in museums are well curated, they are very informative source of information that makes the content-based filtering as a very effective baseline in this domain. In this study, each POI contains 3 museum objects with reach descriptions. In order to build a content-based filtering model, we build a profile of each user after visiting a set of POIs using Language Modeling framework. Each profile's language model is based on all seen objects of \mathbf{p}_{seen} .

Since we have profiles of users at each context, KL-Divergence of each unseen POI's language model and the profile is considered as content-based filtering scores for ranking unseen POIs.

4 POI RECOMMENDATION USING USERS' BEHAVIORS

This section studies how to predict relevant POIs to the given user and context based on users' interaction behaviors, aiming to answer our first research question: *How to understand users' onsite physical behavior and create a behavioral user model that is able to effectively predict relevant unseen POIs?*

In order to model the set-based contextual POI recommendation, we cast the context-aware recommendation problem to a binary classification problem, in which relevant POIs are labeled 1 and irrelevant ones labeled 0. In this way, we try to learn a behavioral model to predict relevant unseen POIs to the given user and context based on the user's interaction behaviors in the context. Then, relevance probability of POIs to the user and context pairs will be used to rank the unseen POIs. To this aim, a set of features that represent users' interaction behaviors in given contexts is defined.

4.1 Feature Set

In order to learn an effective model to rank POIs, we have extracted 18 different features. As it is shown in Table 1, we have classified features to 3 sets, namely, explicit context, onsite and online.

Explicit context refers to information explicitly given by users about the context. In our study, we collected users' gender, their preferred language, their age range and their chosen perspective of the narratives at the exhibition. Previous study on these explicit contexts [16] shows that users behave differently in these different contexts. For example, as it is discussed in [16], children tends to spend less time in front of the POI about the death. Therefore, it seems a reasonable set of features to consider as explicit contexts.

The second group of features is the one gathered onsite without asking users to give any information about their preferences. These features extracted based on users walk-through data. f_5 is the number of seen POIs, which can be considered as a confidence indicator of some other features' scores like f_6 . f_6 is the content-based filtering score of POI candidate based on the profile built using the seen POIs. This content-based filtering score is calculated based on the onsite POI descriptions and users' onsite interactions. That is why it is considered as one of the onsite features in our feature classification.

In addition to f_5 and f_6 , we build users' walk-through graph using their onsite interactions with POIs based on the train set onsite information interaction logs, and calculate $f_7, f_8, f_9, f_{10}, f_{11}$ and f_{12} features. Details of these features are available in Table 1.

The third group of features is defined based on an onsite selected POIs using the onsite users' interactions logs. However, the feature calculation is based on online click-through graph of the museum search engine. Therefore, we classified them as online features. The online click-through graph is filtered to the objects available at onsite POIs. In this study, each onsite POI contains 3 different museum objects. We merge all the objects related to each POI as one node, and the click-through graph's edges are aggregated from all the edges of POIs' objects. As a result, same as onsite walk-through graph, the online click-through graph has onsite POIs as nodes. Details of these features are available in Table 1.

4.2 Learning Model

In order to learn a set-based behavioral POI recommendation model, we have implemented a logistic regression classifier and a deep neural multilayer perceptron with dropouts to estimate relevance of each POI to the given user after visiting a set of POIs. The logistic regression classifier and the deep multilayer perceptron have been trained separately based on each group of features extracted using different users' information interaction behaviors to study which user information interaction behavior is more effective in understanding users' preferences in their interactions with the IoT in smart environment. In the rest of this section, we will detail the logistic regression and the deep multilayer perceptron implemented for the set-based behavioral POI recommendation.

4.2.1 Logistic Regression. Logistic regression classifier is a linear classifier that transparently helps to understand contribution of each feature in estimation of POIs relevancy. In fact, we would like to know which trained logistic classifier performs better and why. To this aim, we train different logistic regression classifiers based on different feature sets using different users' interaction behaviors.

In order to learn a logistic classifier, we use variable $c \in \{0, 1\}$ to show relevance of a POI to a user in a context. Specifically, $P_\theta(c = 1|u, c, p)$ is the relevance score of the POI p to the user u

Table 1: Defined features to predict relevant unseen POIs to users after visiting a set of POIs.

Feature	Category	Description
f_1	Explicit Context	Gender (e.g., Female)
f_2	Explicit Context	Language (e.g., English)
f_3	Explicit Context	Visitor age range (e.g., Adults)
f_4	Explicit Context	Chosen perspective (e.g., Roman)
f_5	Onsite	Seen POIs set size
f_6	Onsite	Content-based relevance score of a POI candidate to a profile created using seen POIs' content that was shown onsite
f_7	Onsite	Unseen POI's PageRank in onsite visits walk-through weighted graph built based on a train set
f_8	Onsite	Unseen POI's PageRank in onsite visits walk-through unweighted graph built based on a train set
f_9	Onsite	Unseen POI's centrality in onsite visits walk-through graph built based on a train set
f_{10}	Onsite	Minimum distance of the seen set of POIs to the POI candidate in the onsite visits walk-through graph built based on a train set
f_{11}	Onsite	Median distance of the seen set of POIs to the POI candidate in the onsite visits walk-through graph built based on a train set
f_{12}	Onsite	Mean distance of the seen set of POIs to the POI candidate in the onsite visits walk-through graph built based on a train set
f_{13}	Online	Unseen POI's PageRank in Online click-through weighted graph built based on a train set
f_{14}	Online	Unseen POI's PageRank in Online click-through unweighted graph built based on a train set
f_{15}	Online	Unseen POI's Centrality in Online click-through graph built based on a train set
f_{16}	Online	Minimum distance of the seen set of POIs to the POI candidate in the Online click-through graph built based on a train set
f_{17}	Online	Median distance of the seen set of POIs to the POI candidate in the Online click-through graph built based on a train set
f_{18}	Online	Mean distance of the seen set of POIs to the POI candidate in the Online click-through graph built based on a train set

and the context c , in which θ is unknown parameters learned using maximum likelihood estimation (MLE) based on the train set. Given the relevance judgments r of each POI p_k to a user u_i and context c_j in the train set, the likelihood L of the train set is as follows:

$$L = \prod_{i=1}^{|U|} \prod_{j=1}^{|C|} \prod_{k=1}^{|P_{seen}|} P_{\theta}(c = 1|u_i, c_j, p_k)^r P_{\theta}(c = 0|u_i, c_j, p_k)^{1-r},$$

in which we assume relevance judgments r are generated independently. We model $P_{\theta}(c = 1|u_i, c_j, p_k)$ by logistic function on a linear combination of features created based on each specific group of users' information interaction behaviors. Then, we optimize the unknown parameters θ by maximizing the following log likelihood function:

$$\begin{aligned} \theta^* = \operatorname{argmax}_{\theta} & \sum_{i=1}^{|U|} \sum_{j=1}^{|C|} \sum_{k=1}^{|P_{seen}|} r \log P_{\theta}(c = 1|u_i, c_j, p_k) \\ & + (1 - r) \log P_{\theta}(c = 0|u_i, c_j, p_k). \end{aligned}$$

In order to turn the logistic classifier scores to probabilities, we have used the softmax function:

$$S(y_i) = \frac{e^{y_i}}{\sum_j e^{y_j}},$$

in which y_i is the logistic classifier score, and $S(y_i)$ is the output relevance probability of our behavioral POI recommendation model.

At last, we rank unseen POIs based on the logistic classifier output probability of POIs' relevancy being estimated based on features created using interaction behaviors of a given user in a context.

4.2.2 Deep Neural Multilayer Perceptron. In this subsection, we investigate on a deep neural multilayer perceptron by an aim of improving effectiveness of the POI recommendation to be used in critical one-shot POI recommendation applications. The motivation behind the critical one-shot POI recommendation is that an irrelevant recommendation sometimes has a very negative effect in users' experience in a way that they might be incorrectly guided to an uninteresting department of a museum that leads to a dissatisfied experience. In this model, for each user in a context, our main goal is to recommend a POI which is highly relevant to them. In the one-shot POI recommendation, we do not care about relevant POIs retrieved after rank 1. In the rest of this section, we detail our deep multilayer perceptron with an aim of improving effectiveness of POI recommendation to be used for the critical one-shot POI recommendation problem.

In order to learn a set based behavioral POI recommendation and learn users' onsite complicated physical behaviors, we have used a deep neural network with 3 hidden layers having 326 units. To learn an effective model and overcome overfitting problem, we have used a dropout feedforward neural network. Let $l \in \{1, 2, 3\}$ be the index of the hidden layers of the network. Let $z^{(l)}$ be the vector of input to layer l and $y^{(l)}$ be the vector of outputs from

layer l . The dropout neural network is modelled as follows for any hidden unit i and $l \in \{0, 1, 2\}$ [19, 28]:

$$\begin{aligned} r^{(l)} &\sim \text{Bernoulli}(p), \\ \tilde{y}^{(l)} &= r^{(l)} * y^{(l)}, \\ z_i^{(l+1)} &= w_i^{(l+1)} \tilde{y}^{(l)} + b_i^{(l+1)}, \\ y_i^{(l+1)} &= f(z_i^{(l+1)}), \end{aligned}$$

where $r^{(l)}$ denotes a vector of independent Bernoulli random variables having probability p of being 1, $\tilde{y}^{(l)}$ is thinned outputs created by multiplying a sample of $r^{(l)}$ vector by outputs of layer l (i.e., $y^{(l)}$) and used as input for the next layer $l + 1$, $w^{(l)}$ and $b^{(l)}$ are weights and biases at layer l , and f is an activation function, which is rectified linear units (ReLUs) in our setup. This process is done at each layer.

Following many researches in neural network domain, we have used $p = 0.5$ in our dropout network. This value is reported as a close to optimal value for a wide range of networks in different applications [28].

In the learning phase, the derivatives of the loss function are back-propagated through the dropout network. The dropout network is trained using the stochastic gradient descent (SGD) algorithm with mini batches, which is widely used algorithm for training neural networks. The learning rates are adjusted based on adaptive gradient algorithm (AdaGrad) [8]. In the test phase, the sub-network is used without dropout, but the weights are scaled as $W_{test}^{(l)} = pW^{(l)}$.

For the classification purpose and having probabilities as outputs, we have used Logistic classifier in the last layer. The logistic classifier in the last layer is trained same as the logistic regression classifier being discussed in previous subsection. The only difference is that, in the logistic classifier being used in the last layer, we model $P_\theta(c = 1|u_i, c_j, p_k)$ by logistic function on a linear combination of inputs from the last hidden layer units' outputs. At last, the final relevance probability of $P_\theta(c = 1|u_i, c_j, p_k)$ is used to rank unseen POIs based on features created using interaction behaviors of a given user in a context.

5 EXPERIMENTAL SETUP

In this section, we describe our experimental setup. We first describe the data set used in this paper, and second detail the evaluation methodology used in this study.

5.1 Dataset

The dataset of this study is based on onsite and online interaction logs collected at an archeological museum. In this archeological museum, RFID tags are provided as a key to access some additional information about objects being shown in the museum. Users can enter their preferences at the beginning of the museum exhibition in order to personalize the content being shown in all of the POIs. These preferences are perspectives of the narratives, language, gender and the user's age range.

After checking in, users are free to put their tags on RFID readers of POIs to unlock contents being shown about objects at the POIs. Each POI contains 3 different archeological objects. Users are free

Table 2: An example of records created for the test collection using a user session. The judgments are based on seen POI set-size 2 and 3.

Context	Query Seen POI set	Candidate	Relevance
c_1	<POI1,POI2>	POI3	0
c_1	<POI1,POI2>	POI4	1
c_1	<POI1,POI2>	POI5	0
c_1	<POI1,POI2>	POI6	0
c_1	<POI1,POI2>	POI7	1
c_1	<POI1,POI2>	POI8	0
c_1	<POI1,POI2,POI4>	POI3	0
c_1	<POI1,POI2,POI4>	POI5	0
c_1	<POI1,POI2,POI4>	POI6	0
c_1	<POI1,POI2,POI4>	POI7	1
c_1	<POI1,POI2,POI4>	POI8	0

to interact with POIs in any order. They can watch short movies, interact with 3D photos of POIs' objects, or read contents about objects being shown at POIs. At last, users might check out in a summary station, in which they might leave their name, birth date and email. In this paper, 5 months onsite logs of the museum with more than 21,000 sessions is used, which leads to 3,925 high-quality onsite sessions to be used for evaluation purposes.

In addition to the users' onsite information interaction logs, we also collected query logs of the museum search engine. The online features, detailed in Table 1, have been extracted based on 18,001 high-quality sessions created by filtering bot sessions.

5.2 Evaluation Methodology

In our collected onsite information interaction logs, about 16,000 out of 21,000 sessions either did not have any interactions with POIs or they did not check out at the summary station, and about 1,000 of them had interactions with all the POIs. In order to avoid bias over users who are interested in visiting all or none of the POIs at the museum, we exclude all sessions have checked in at all or none of the POIs at the exhibition. As a result of this preprocessing step, 3,925 out of 21,000 high-quality onsite information interaction sessions remains for creating the test collection.

Considering the walk-through graph, for each user in a session and at each checked in POI during their visit, we created a test collection using the seen set of POIs, the user and the explicit contexts as the query and the unseen POIs as the candidates, for which we have judgments based on the user's session. Basically, we know which POI candidates are visited by the user and consider them as relevant POIs. The rest of the POIs are considered as irrelevant POIs.

Doing the above procedure in building the test collection leads to create a contextual set-based POI recommendation test collection having 1,083,623 judgments. Table 2 shows an example of records created using a user session. To test our proposed model, in order to avoid overfitting, we have done 5-fold cross-validation, in which for each fold as a test set, 3 out of the 4 remained folds randomly sampled and used as a train set, and the remained fold used as a

validation set. We repeat the process for all the five folds and report the average of the evaluation metrics.

5.3 Evaluation Metrics

For the evaluation of the defined set-based behavioral POI recommendation task, we cast the problem to a ranking task and use mean reciprocal-rank (*MRR*), mean average precision (*MAP*) and *R*-precision (*R-Prec*) as metrics that are effective to evaluate proposed models. Moreover, in order to evaluate the one-shot POI recommendation systems, we use precision at rank 1 (*P@1*) as an evaluation metric.

The *MRR* is the average of the reciprocal ranks of the first relevant result for a set of queries Q as $MRR = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{rank_i}$. For a single query, *AP* is defined as the average of the $p@n$ values (i.e., $p@n = \frac{\# \text{ relevant POIs in top } n \text{ results}}{n}$) for all relevant POIs as $AP = \frac{\sum_{n=1}^N p@n \times rel(n)}{R}$, in which n is the rank, N is the number of retrieved POIs candidates, and $rel(n)$ is a binary function indicating the relevance of a given rank. *MAP* is the mean value of the *AP*s computed for all queries. *R-Prec* is precision at rank R where R is the number of relevant candidates for the given query. At last, $p@1$ is the precision at rank 1.

6 EXPERIMENTAL RESULTS

In this section, we provide answer to the research questions stated in the introduction section.

6.1 POI Recommendation Using Users' Information Interaction Behaviors

This section answer our second research question: *How strong are different users' interaction behaviors with IoT in understanding users' preferences?*

To this aim, we have used each of the three groups of features extracted based on each information interaction behaviors to train a POI recommendation system. Specifically, we have trained three different logistic regression classifiers, which are trained based on: 1) the explicit context features (i.e., Logistic Regression-Explicit Context) 2) the onsite features (i.e., Logistic Regression-Onsite) and 3) the online features (i.e., Logistic Regression-Online).

In the rest of this subsection, we first investigate whether users' online digital interaction behaviors are similar to the users' onsite physical behavior. Then, we detail relative importance of each feature extracted based on features' weights being learned by logistic regression classifiers using each type of users' interaction behaviors with an aim of understanding users' behaviors.

6.1.1 Onsite Physical Behavior vs. Online Digital Behavior. We first look at the question: *Are online digital behaviors similar to onsite physical behaviors? Does understanding online digital users' information interaction behaviors have a positive effect in learning a model to predict unseen relevant POIs and complete users' personalized onsite visits?*

In order to answer this research question, we compare POI recommendation systems trained based on each type of interaction behavior. As it is shown in Figure 3, the POI recommendation system trained based on users' online digital interaction behavior is not only as good as the other POI recommendation systems being

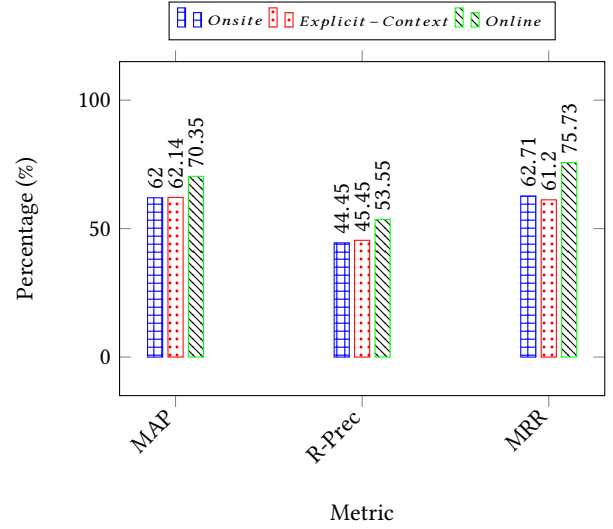


Figure 3: Effectiveness of different types of users' interaction behavior in understanding their onsite preferences.

trained based on either explicit context or onsite interaction behaviors, but also is performing better than them in terms of all common tested information retrieval metrics. This experiment indicates that availability of the considerable amount of online interaction logs in comparison to onsite interaction logs leads to training an effective onsite POI recommendation system based on users' online digital interaction behaviors. As we achieve an effective onsite POI recommendation system based on users' online digital interaction behaviors, we conclude that there is a similarity between onsite physical and online digital information interaction behaviors.

6.1.2 Features Relative Importance in Understanding Users' Interaction Behaviors. We now look at the question: *What are the relative importance of each feature extracted based on different users' interaction behaviors in effectiveness of POI recommendation systems?*

To this aim, we normalize features' weights being learned in each logistic regression classifier trained for each group of features separately. Then, average of the normalized features' weights over the 5-fold cross-validation are reported and compared in Figure 4.

As it is shown in Figure 4, among the explicit context interaction, the chosen language (i.e., f_2) at the start of museum visits is relatively more important in comparison to other explicit context based features. Moreover, mean distance of the seen POIs to a POI candidate in the onsite visits' walk-through graph (i.e., f_{12}) has relatively more importance in comparison to other onsite interaction behavior based features. Regarding the online interaction behaviors, median distance of the seen set of POIs to the given candidate in the online click-through graph (i.e., f_{17}) is relatively more important than other online features in the effectiveness of the POI recommendation systems.

6.2 One-Shot POI Recommendation Using Users' Interaction Behaviors

This section answer our third research question: *How effective is behavioral POI recommendation system in one-shot POI recommendation problem?* To this aim, we study effectiveness of the

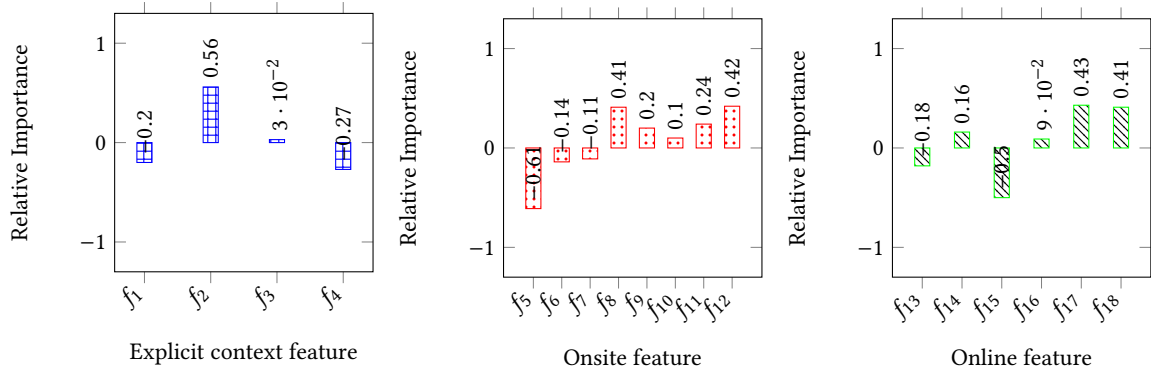


Figure 4: Features’ relative importance in POI recommendation trained based on each group of users’ information interaction behaviors.

Table 3: Set-based one-shot POI recommendation effectiveness comparison between the *Deep MLP-Online* and baselines. * indicates the improvement is statistically significant ($\rho < 0.05$).

Run	P@1	MRR
<i>Content-based Filtering</i>	57.45	75.68
<i>Popularity</i>	60.86	77.67
<i>Bias-Based Filtering</i>	61.57	77.71
<i>Logistic Regression-Online</i>	56.97	75.73
<i>Deep MLP-Online</i>	75.81 (23.12%*)	86.39 (11.17%*)

implemented deep multilayer perceptron in one-shot onsite POI recommendation problem in comparison to the discussed baselines as well as the logistic regression POI recommendation system. In addition to baselines’ effectiveness, Table 3 shows performance of the best deep multilayer perceptron (i.e., Deep MLP) and logistic regression classifiers, trained based on online digital interaction behaviors, in terms of $p@1$ and MRR .

As it is shown in Table 3, the deep MLP significantly improves the best hard-to-beat baseline (i.e., Bias-Based Filtering) in one-shot POI recommendation. In particular, the deep MLP has 23.12% improvement over the bias-based filtering baseline in term of $p@1$, which is the only used metric that measure one-shot POI recommendation performance. This experimental result shows that our proposed deep MLP one-shot POI recommendation system is very effective, and can lead to an interesting personalized experience in such a critical application.

7 DISCUSSION AND CONCLUSIONS

The main focus of this paper is the study of how to build a behavioral user model for the set-based POI recommendation problem using users’ both onsite and online information interaction behaviors. A study on the strength of using each type of users’ interaction behaviors with IoT in understanding users’ onsite information interaction preferences shows that POI recommendation systems trained using features extracted from a combination of both onsite physical and online digital information interaction behaviors (i.e.,

online features) performs better than the ones trained by explicitly given context or onsite information interaction behavior. Therefore, we conclude that there is a similarity between onsite physical and online digital interaction preferences that causes an improvement on the onsite POI recommendation effectiveness. Moreover, we have studied the critical one-shot POI recommendation problem. According to our analysis, the learned models based on just basic explicit given contexts or onsite users’ behaviors do not improve the hard-to-beat defined baselines (i.e., popularity and bias-based filtering). However, using a deep multilayer perceptron based on features extracted by online interaction behaviors leads to a significant improvement over the best baseline in all the defined evaluation metrics. Specifically, it has a statistically significant improvement over all baselines with 23% improvement in term of $p@1$ and 11% improvement in term of MRR . Therefore, our proposed approach is very effective in critical one-shot POI recommendation. The proposed behavioral user model is very general and can be widely used in any environment with an integrated Internet of Things (IoT) infrastructure. Specifically, in the Cultural Heritage domain, the implemented technology being used in this study and implemented within the European meSch project¹, is listed in [9, 20] as one of the implemented technologies in museums that provides a more interactive and multisensory experiences for visitors. This is one of the technologies mentioned in museum edition of the NMC Horizon 2015 and 2016 reports as a technology being integrated in museums in four to five years time-to-adoption horizon. As a future work, we are going to increase number of POIs in the museum and see how effective is the proposed behavioral set-based POI recommendation model for the bigger datasets. Moreover, in addition to the evaluation detailed in this paper based on a high volume of real users, we are eager to do a user study to test our proposed behavioral user model. As another line of future work, we will investigate on using recursive neural network to improve our proposed behavioral user models to be used in contextual suggestion problem [15].

ACKNOWLEDGMENTS

This research is funded in part by the European Community’s FP7 (project meSch, grant # 600851).

¹<http://www.mesch-project.eu/>

REFERENCES

- [1] Gediminas Adomavicius and Alexander Tuzhilin. 2011. Context-aware recommender systems. In *Recommender systems handbook*. Springer, 217–253.
- [2] Luigi Atzori, Antonio Iera, and Giacomo Morabito. 2010. The internet of things: A survey. *Computer networks* 54, 15 (2010), 2787–2805.
- [3] Payam Barnaghi, Wei Wang, Cory Henson, and Kerry Taylor. 2012. Semantics for the Internet of Things: early progress and back to the future. *International Journal on Semantic Web and Information Systems (IJSWIS)* 8, 1 (2012), 1–21.
- [4] Ilaria Bartolini, Vincenzo Moscato, Ruggero G Pensa, Antonio Penta, Antonio Picariello, Carlo Sansone, and Maria Luisa Sapino. 2016. Recommending multimedia visiting paths in cultural heritage applications. *Multimedia Tools and Applications* 75, 7 (2016), 3813–3842.
- [5] Betim Berjani and Thorsten Strufe. 2011. A Recommendation System for Spots in Location-based Online Social Networks. In *Proceedings of the 4th Workshop on Social Network Systems (SNS '11)*. ACM, 4:1–4:6.
- [6] Fabian Bohnert, Ingrid Zukerman, and Junaidy Laures. 2012. GECKOmmender: Personalised theme and tour recommendations for museums. In *International Conference on User Modeling, Adaptation, and Personalization*. Springer, 26–37.
- [7] Aleksandr Chuklin, Ilya Markov, and Maarten de Rijke. 2015. Click Models for Web Search. *Synthesis Lectures on Information Concepts, Retrieval, and Services* 7, 3 (2015), 1–115.
- [8] John Duchi, Elad Hazan, and Yoram Singer. 2011. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research* 12, Jul (2011), 2121–2159.
- [9] A Freeman, S Adams Becker, M Cummins, E McKelroy, C Giesinger, and B Yuhnke. 2016. NMC horizon report: 2016: Museum edition. *The New Media Consortium, Austin, Texas 2016* (2016).
- [10] Peter Friess. 2013. *Internet of things: converging technologies for smart environments and integrated ecosystems*. River Publishers.
- [11] Huiji Gao, Jiliang Tang, Xia Hu, and Huan Liu. 2013. Exploring Temporal Effects for Location Recommendation on Location-based Social Networks. In *Proceedings of the 7th ACM Conference on Recommender Systems (RecSys '13)*. ACM, 93–100.
- [12] Karl Grieser, Timothy Baldwin, and Steven Bird. 2007. Dynamic path prediction and recommendation in a museum environment. In *Proc. of the Workshop on Language Technology for Cultural Heritage Data (LaT-eCH 2007)*. 49–56.
- [13] Jean-Benoit Griesner, Tael Abdessalem, and Hubert Naacke. 2015. POI Recommendation: Towards Fused Matrix Factorization with Geographical and Temporal Influences. In *Proceedings of the 9th ACM Conference on Recommender Systems (RecSys '15)*. ACM, 301–304.
- [14] Ido Guy. 2015. The Role of User Location in Personalized Search and Recommendation. In *Proceedings of the 9th ACM Conference on Recommender Systems (RecSys '15)*. ACM, 236–236.
- [15] Seyyed Hadi Hashemi, Charles L. A. Clarke, Jaap Kamps, Julia Kiseleva, and Ellen M. Voorhees. 2016. Overview of the TREC 2016 Contextual Suggestion Track. In *Proceeding of Text REtrieval Conference (TREC)*.
- [16] Seyyed Hadi Hashemi, Wim Hupperetz, Jaap Kamps, and Merel van der Vaart. 2016. Effects of Position and Time Bias on Understanding Onsite Users' Behavior. In *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval (CHIIR '16)*. ACM, 277–280.
- [17] Seyyed Hadi Hashemi and Jaap Kamps. 2017. Skip or Stay: Users' Behavior in Dealing with Onsite Information Interaction Crowd-Bias. In *Proceedings of the 2017 ACM on Conference on Human Information Interaction and Retrieval (CHIIR '17)*. ACM.
- [18] José M Hernández-Muñoz, Jesús Bernat Vercher, Luis Muñoz, José A Galache, Mirko Presser, Luis A Hernández Gómez, and Jan Pettersson. 2011. Smart cities at the forefront of the future internet. In *The Future Internet Assembly*. Springer, 447–462.
- [19] Geoffrey E. Hinton, Nitish Srivastava, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2012. Improving neural networks by preventing co-adaptation of feature detectors. *CoRR abs/1207.0580* (2012).
- [20] Larry Johnson, Samantha Adams Becker, Victoria Estrada, and Alex Freeman. 2015. *The NMC Horizon Report: 2015 Museum Edition*. ERIC.
- [21] Alexandros Karatzoglou, Xavier Amatriain, Linas Baltrunas, and Nuria Oliver. 2010. Multiverse Recommendation: N-dimensional Tensor Factorization for Context-aware Collaborative Filtering. In *Proceedings of the Fourth ACM Conference on Recommender Systems (RecSys '10)*. ACM, 79–86.
- [22] Xin Liu and Karl Aberer. 2013. SoCo: A Social Network Aided Context-aware Recommender System. In *Proceedings of the 22Nd International Conference on World Wide Web (WWW '13)*. ACM, 781–802.
- [23] Claudio Lucchese, Raffaele Perego, Fabrizio Silvestri, Hossein Vahabi, and Rossano Venturini. 2012. How Random Walks Can Help Tourism. In *Proceedings of Advances in Information Retrieval: 34th European Conference on IR Research, ECIR 2012*.
- [24] Moon-Hee Park, Jin-Hyuk Hong, and Sung-Bae Cho. 2007. Location-based recommendation system using bayesian user's preference model in mobile devices. In *International Conference on Ubiquitous Intelligence and Computing*. Springer, 1130–1139.
- [25] Charith Perera, Arkady Zaslavsky, Peter Christen, and Dimitrios Georgakopoulos. 2014. Context aware computing for the internet of things: A survey. *IEEE Communications Surveys & Tutorials* 16, 1 (2014), 414–454.
- [26] Steffen Rendle, Zeno Gantner, Christoph Freudenthaler, and Lars Schmidt-Thieme. 2011. Fast Context-aware Recommendations with Factorization Machines. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '11)*. ACM, 635–644.
- [27] Christoph Scholz, Jens Illig, Martin Atzmueller, and Gerd Stumme. 2014. On the Predictability of Talk Attendance at Academic Conferences. In *Proceedings of the 25th ACM Conference on Hypertext and Social Media*. ACM, 279–284.
- [28] Nitish Srivastava, Geoffrey E Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research* 15, 1 (2014), 1929–1958.
- [29] Gang Wang, Xinyi Zhang, Shiliang Tang, Haitao Zheng, and Ben Y. Zhao. 2016. Unsupervised Clickstream Clustering for User Behavior Analysis. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, 225–236.
- [30] Xiangye Xiao, Yu Zheng, Qiong Luo, and Xing Xie. 2010. Finding Similar Users Using Category-based Location History. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems (GIS '10)*. ACM, 442–445.
- [31] Mao Ye, Peifeng Yin, Wang-Chien Lee, and Dik-Lun Lee. 2011. Exploiting Geographical Influence for Collaborative Point-of-interest Recommendation. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '11)*. ACM, 325–334.
- [32] Josh Jia-Ching Ying, Eric Hsueh-Chan Lu, Wen-Ning Kuo, and Vincent S. Tseng. 2012. Urban Point-of-interest Recommendation by Mining User Check-in Behaviors. In *Proceedings of the ACM SIGKDD International Workshop on Urban Computing (UrbComp '12)*. ACM, 63–70.
- [33] Vincent W. Zheng, Bin Cao, Yu Zheng, Xing Xie, and Qiang Yang. 2010. Collaborative Filtering Meets Mobile Recommendation: A User-centered Approach. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI'10)*. AAAI Press, 236–241.
- [34] Vincent W. Zheng, Yu Zheng, Xing Xie, and Qiang Yang. 2010. Collaborative Location and Activity Recommendations with GPS History Data. In *Proceedings of the 19th International Conference on World Wide Web (WWW '10)*. ACM, 1029–1038.
- [35] Vincent W. Zheng, Yu Zheng, Xing Xie, and Qiang Yang. 2010. Collaborative Location and Activity Recommendations with GPS History Data. In *Proceedings of the 19th International Conference on World Wide Web (WWW '10)*. ACM, 1029–1038.
- [36] Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. 2009. Mining Interesting Locations and Travel Sequences from GPS Trajectories. In *Proceedings of the 18th International Conference on World Wide Web (WWW '09)*. ACM, 791–800.
- [37] Jinfeng Zhuang, Tao Mei, Steven C.H. Hoi, Ying-Qing Xu, and Shipeng Li. 2011. When Recommendation Meets Mobile: Contextual and Personalized Recommendation on the Go. In *Proceedings of the 13th International Conference on Ubiquitous Computing (UbiComp '11)*. ACM, 153–162.
- [38] Onno Zoeter. 2015. Recommendations in Travel. In *Proceedings of the 9th ACM Conference on Recommender Systems (RecSys '15)*. ACM, 234–234.